

**Identification and heterologous reconstitution of a 5-alk(en)ylresorcinol synthase  
from endophytic fungus *Shiraia* sp. Slf14**

Huiwen Yan<sup>1,2,†</sup>, Lei Sun<sup>1,†</sup>, Jinge Huang<sup>1</sup>, Yixing Qiu<sup>1,3</sup>, Fuchao Xu<sup>1</sup>, Riming Yan<sup>1,4</sup>,  
Du Zhu<sup>4</sup>, Wei Wang<sup>3,\*</sup>, and Jixun Zhan<sup>1,3,\*</sup>

<sup>1</sup> Department of Biological Engineering, Utah State University, 4105 Old Main Hill,  
Logan, UT 84322-4105, USA

<sup>2</sup> The State Key Laboratory of Medical Genetics & School of Life Sciences, Central  
South University, Changsha, Hunan 410083, China

<sup>3</sup> TCM and Ethnomedicine Innovation & Development Laboratory, School of Pharmacy,  
Hunan University of Chinese Medicine, Changsha, Hunan 410208, China

<sup>4</sup> Key Laboratory of Protection and Utilization of Subtropic Plant Resources of Jiangxi  
Province, College of Life Science, Jiangxi Normal University, Nanchang, Jiangxi  
330022, China

<sup>†</sup> These authors contributed equally to this work.

\* For correspondence. E-mails: [jixun.zhan@usu.edu](mailto:jixun.zhan@usu.edu) (J.Z.);  
[wangwei402@hotmail.com](mailto:wangwei402@hotmail.com) (W.W.). Tel.: +1 435 797 8774.

Short title: A 5-alk(en)ylresorcinol synthase from endophytic fungus *Shiraia* sp. Slf14

## Abstract

A new type III polyketide synthase gene (*Ssars*) was discovered from the genome of *Shiraia* sp. Slf14, an endophytic fungal strain from *Huperzia serrata*. The intron-free gene was cloned from the cDNA and ligated to two expression vectors pET28a and YEpADH2p-URA3 for expression in *Escherichia coli* BL21(DE3) and *Saccharomyces cerevisiae* BJ5464, respectively. SsARS was efficiently expressed in *E. coli* BL21(DE3), leading to the synthesis of a series of polyketide products. Six major products were isolated from the engineered *E. coli* and characterized as 1,3-dihydroxyphenyl-5-undecane, 1,3-dihydroxyphenyl-5-*cis*-6'-tridecene, 1,3-dihydroxyphenyl-5-tridecane, 1,3-dihydroxyphenyl-5-*cis*-8'-pentadecene, 1,3-dihydroxyphenyl-5-pentadecane and 1,3-dihydroxyphenyl-5-*cis*-10'-heptadecene, respectively, based on the spectral data and biosynthetic origin. Expression of SsARS in the yeast also led to the synthesis of the same polyketide products, indicating that this enzyme can be reconstituted in both heterologous hosts. Supplementation of soybean oil into the culture of *E. coli* BL21(DE3)/SsARS increased the production titers of **1-6** and led to the synthesis of an additional product, which was identified as 5-(8'Z,11'Z-heptadecadienyl)resorcinol. This work thus allowed the identification of SsARS as a 5-alk(en)ylresorcinol synthase with flexible substrate specificity toward endogenous and exogenous fatty acids. Desired resorcinol derivatives may be synthesized by supplying corresponding fatty acids into the culture medium.

**Keywords** *Shiraia* sp. Slf14, type III polyketide synthase, akyresorcinols, substrate flexibility, precursor-directed biosynthesis

## Introduction

Fungi are known to produce a large variety of bioactive natural products, such as lovastatin (anti-cholesterol) and penicillin (antibacterial). Polyketides are a large family of structurally diverse natural products including erythromycin (antibacterial), tetracycline (antibacterial) and doxorubicin (antitumor). They are assembled from acyl-coenzyme A precursors by polyketide synthases (PKSs), which are classified into three groups based on sequences and functional mechanisms, including types I, II, and III. Type I PKS, similar to type I fatty acid synthase (FAS), is a giant assembly of multifunctional polypeptides, each consisting of a series of catalytic domains (Keatinge-Clay 2012). Representative bioactive compounds synthesized by type I PKSs are lovastatin and erythromycin. Type II PKSs are discrete enzymes that work collaboratively to form aromatic polyketide products, such as tetracycline and doxorubicin (Zhan 2009). Type III PKSs are self-contained enzymes that form homodimers. These single enzymes can catalyze precursor priming, chain extension and cyclization. Type III PKSs are involved in the biosynthesis of a wide array of natural products such as chalcones, pyrones, acridones, phloroglucinols, stilbenes, and resorcinolic lipids (Yu et al. 2012). All three types of PKS are present in bacteria and type II PKSs are exclusively found in bacteria such as actinomycetes. Plants contain various type III PKSs, while fungi are known to harbor type I and type III PKSs. These PKSs are involved in the formation of various metabolites in these hosts.

Alk(en)ylresorcinols are phenolic lipids synthesized by type III PKSs. They are commonly present in several plant families and have been proposed as specific dietary

biomarkers of whole-grain wheat and rye intake (Kyro et al. 2014; Landberg et al. 2008). They were also found in fungi, algae, mosses, bacteria, and marine sponges (Landberg et al. 2014). Many of natural alk(en)ylresorcinols have shown antimicrobial activities. For example, DB-2017 is an alkylresorcinol antibiotic from *Pseudomonas* sp. B-9094. It has shown antimicrobial activity against mycobacteria, Gram-positive bacteria, and fungi (Kanda et al. 1975; Kitahara and Kanda 1975). Some resorcinol derivatives were also found to possess cytotoxic, anticarcinogenic, antiproliferative, antileishmanial and antioxidant properties (Luis et al. 2016).

Symbiotic microorganisms in plants live closely with the hosts and are found to produce various valuable compounds. *Shiraia* sp. Slf14 is a promising endophytic fungus from *Huperzia serrata*. This strain possesses the ability to produce hypocrellins and huperzine A (Yang et al. 2014). As polyketide natural products, hypocrellins are a group of molecules with a good prospect of clinical applications, and have been used as an effective photosensitizer for photodynamic therapy (PDT) of tumors (Ali and Olivo 2002; Ali et al. 2001; Estey et al. 1996). Huperzine A is a sesquiterpene alkaloid originally isolated from *H. serrata*. It acts as a cholinesterase inhibitor and thus has potential for use in the treatment of Alzheimer's disease (Damar et al. 2017).

To better understand *Shiraia* sp. Slf14, we have sequenced the genome of this fungus (Yang et al. 2014). Analysis of the sequence genome revealed a putative type III PKS. Here we report the identification of this type III PKS and heterologous reconstitution of this enzyme in *Escherichia coli* and *Saccharomyces cerevisiae*. Six alk(en)ylresorcinols were synthesized by this enzyme from malonyl-CoA and various

fatty acyl starter units. An additional alkylresorcinol was generated by *E. coli* when soybean oil was supplemented into the medium.

## **Materials and Methods**

### **Strains and media**

DNA manipulations were performed using standard techniques with *E. coli* XL1-Blue as a host organism. Expression vector pET28a was purchased from Novagen (Madison, WI, USA). *E. coli* BL21(DE3) purchased from Stratagene (La Jolla, CA, USA) was used as the expression host. The *E. coli* strains were routinely grown in LB medium. *Shiraia sp.* Slf14 (CCTCC No. M209294) was previously isolated from *H. serrata* and was grown in potato dextrose broth (PDB) medium (Zhu et al. 2010). *S. cerevisiae* BJ5464 (ATCC 208288) was obtained from the American Type Culture Collection and routinely grown in YPD medium.

### **Construction of pET28a-*Ssars* (pHW4) and YEpADH2p-URA3-*Ssars* (pHW8)**

Wild type *Shiraia sp.* Slf14 was grown in PDB for 3 days. The cells were harvested by filtration, from which the RNA was extracted with TRIzol Reagent (Invitrogen, USA). The extracted RNA was purified with a Direct-zol RNA MiniPrep kit (Zymo Research, CA, USA). The cDNA was synthesized from mRNA with M-MuLV Reverse Transcriptase (NEB, MA, USA). With the cDNA as a template, the 1,278-bp DNA intron-free *Ssars* gene (GenBank accession no. MH048853) was amplified by PCR with a set of specific primers including SsARS-F-NheI (5'-aaGCTAGCcatgtcccacgtggctaacac-3', with a *NheI* site shown by underlining) and SsARS-R-BamHI (5'-aaGGATCCttaacctgttcgtacagtgt-3', with an *BamHI* site shown by

underlining). The resulting PCR product was cloned into the pJET1.2 cloning vector. After sequencing, this gene subsequently cloned into the pET28a vector between the *NheI* and *BamHI* sites, yielding pHW4. The same gene was amplified by PCR using another primer set including the above mentioned SsARS-F-*NheI* and SsARS-R-*PmeI* (5'-aaGTTTAAACttaacctgttcgtacagtgt-3', with a *PmeI* site shown by underlining). The resulting PCR product was cloned into pJET1.2, and subsequently inserted into the yeast expression vector YEpADH2p-URA3 between the *NheI* and *PmeI* sites, yielding pHW8.

#### **SDS-PAGE analysis of the expression of SsARS in *E. coli* BL21(DE3)**

*E. coli* BL21(DE3)/pHW4 was grown in 50 mL of LB medium supplemented with 50 µg/mL kanamycin at 37°C to a certain OD<sub>600</sub> 0.4-0.6. The culture was then induced by 200 µM isopropyl-1-thio-β-D-galactoside (IPTG) and incubated at 28°C and 250 rpm for 12 hrs. Cells were harvested by centrifugation at 4,000 rpm for 10 min, resuspended in the lysis buffer composed of 20 mM Tris-HCl (pH 7.9) and 0.5 M NaCl, and disrupted by sonication. A crude cell lysate was prepared by removal of cell debris by centrifugation at 13,000 rpm at 4 °C for 15 min, which was then analyzed by SDS-PAGE.

#### **LC-MS analysis of the polyketide products in *E. coli* BL21(DE3)/pHW4**

For production analysis, *E. coli* BL21(DE3)/pHW4 was grown and induced as described above. The pellet was extracted with methanol and the supernatant with an equal volume of ethyl acetate three times. The extracts were combined and analyzed on an Agilent 1200 HPLC instrument (Agilent Eclipse Plus C18, 5 µm, 4.6 mm × 250 mm)

connected with an Agilent 6130 Single Quadrupole mass spectrometer, eluted with a linear gradient of 75 to 90 % (v/v) acetonitrile-water (containing 0.1% trifluoroacetic acid) over 45 min at a flow rate of 1 mL/min. The products were detected at 210 nm.

#### **Purification and structural characterization of alk(en)ylresorcinols 1–6**

To isolate compounds **1–6** for structural characterization, the culture of *E. coli* BL21(DE3)/pHW4 was scaled up to 200 mL. Briefly, 200  $\mu$ L of seed culture (OD<sub>600</sub>=1.0) of *E. coli* BL21(DE3)/pHW4 was inoculated into 200 mL of LB medium containing 50  $\mu$ g/mL kanamycin, and grown to OD<sub>600</sub> of 1.0. 200  $\mu$ M IPTG was then added to induce protein expression. After an additional 72 hrs, the culture was extracted as described above. To isolate these compounds, a typical purification procedure was given below. The crude extract was fractionated on a silica gel-60 column, eluted with a stepwise gradient of ethyl acetate-hexane (0:100, 1:99, 2:98, 3:97, 5:95, 10:90, 15:85, 20:80 v/v, each 250 mL) to afford 8 fractions. The fractions containing the target compounds were further separated by reverse-phased HPLC (Agilent Eclipse Plus C18, 5  $\mu$ m, 4.6 $\times$ 250 mm), eluted with acetonitrile-water at a flow rate of 1 mL/min. Compounds **1** and **2** were found to be in fraction 6 after silica gel column chromatography. This fraction was further separated on HPLC with 67% acetonitrile-water. The peaks at 20.5 min and 23.7 min were collected to yield 1.7 mg of **1** and 1.4 mg of **2** in pure form. Compounds **3** and **4** were found to be in fraction 5 after silica gel column chromatography. This fraction was further separated on HPLC with 74% acetonitrile-water. The peaks at 32.4 min and 35.6 min were collected to yield 1.5 mg of **3** and 2.1 mg of **4**, respectively, in pure form. Compounds **5** and **6** were found to be

in fraction 3 after silica gel column chromatography. This fraction was further separated on HPLC with 82% acetonitrile-water. The peaks at 47.9 min and 50.5 min were collected to yield 0.9 mg of **5** and 1.2 mg of **6**, respectively. Electrospray ionization mass spectrometry (ESI-MS) spectra were collected on an Agilent 6130 Single quadrupole LC-MS. Nuclear magnetic resonance (NMR) spectra were acquired in CDCl<sub>3</sub> on a Bruker NMR instrument (500 MHz for <sup>1</sup>H NMR). The chemical shift ( $\delta$ ) values were given in parts per million (ppm). The coupling constants ( $J$  values) were reported in Hertz (Hz).

Compound **1**: ESI-MS (+): [M+H]<sup>+</sup>  $m/z$  265.2; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.26 (2H, d,  $J$  = 1.8 Hz, H-4 and H-6), 6.22 (1H, t,  $J$  = 1.8 Hz, H-2), 2.49 (2H, t,  $J$  = 7.5 Hz, H-1'), 1.58 (2H, m, H-2'), 1.27 (approx. 16H, brs), 0.90 (3H, t,  $J$  = 6.9 Hz, H-11').

Compound **2**: ESI-MS (+): [M+H]<sup>+</sup>  $m/z$  291.2; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.26 (2H, d,  $J$  = 1.8 Hz, H-4 and H-6), 6.20 (1H, t,  $J$  = 1.8 Hz, H-2), 5.38-5.36 (2H, m, H-6' and H-7'), 2.51 (2H, t,  $J$  = 7.5 Hz, H-1'), 2.02 (4H, m, H-5' and H-8'), 1.59 (2H, m, H-2'), 1.34 (approx. 12H, brs), 0.92 (3H, t,  $J$  = 6.9 Hz, H-13').

Compound **3**: ESI-MS (+): [M+H]<sup>+</sup>  $m/z$  293.2; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.26 (2H, d,  $J$  = 1.8 Hz, H-4 and H-6), 6.20 (1H, t,  $J$  = 1.8 Hz, H-2), 2.49 (2H, t,  $J$  = 7.5 Hz, H-1'), 1.59 (2H, m, H-2'), 1.32 (approx. 20H, brs), 0.91 (3H, t,  $J$  = 6.9 Hz, H-13').

Compound **4**: ESI-MS (+): [M+H]<sup>+</sup>  $m/z$  319.2; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>):  $\delta$  6.27 (2H, d,  $J$  = 1.8 Hz, H-4 and H-6), 6.21 (1H, t,  $J$  = 1.8 Hz, H-2), 5.38-5.36 (2H, m, H-8' and H-9'), 2.49 (2H, t,  $J$  = 7.5 Hz, H-1'), 2.03 (4H, m, H-7' and H-10'), 1.59 (2H, m, H-2'), 1.29 (approx. 16H, brs), 0.89 (3H, t,  $J$  = 6.9 Hz, H-15').



Compound **5**: ESI-MS (+):  $[M+H]^+$   $m/z$  321.4;  $^1H$  NMR (500 MHz,  $CDCl_3$ ):  $\delta$  6.26 (2H, d,  $J = 1.8$  Hz, H-4 and H-6), 6.19 (1H, t,  $J = 1.8$  Hz, H-2), 2.51 (2H, t,  $J = 7.5$  Hz, H-1'), 1.59 (2H, m, H-2'), 1.28 (approx. 24H, brs), 0.90 (3H, t,  $J = 6.9$  Hz, H-15').

Compound **6**: ESI-MS (+):  $[M+H]^+$   $m/z$  347.3;  $^1H$  NMR (500 MHz,  $CDCl_3$ ):  $\delta$  6.26 (2H, d,  $J = 1.8$  Hz, H-4 and H-6), 6.20 (1H, t,  $J = 1.8$  Hz, H-2), 5.37-5.36 (2H, m, H-10' and H-11'), 2.50 (2H, t,  $J = 7.5$  Hz, H-1'), 2.03 (4H, m, H-9' and H-12'), 1.57 (2H, m, H-2'), 1.34 (approx. 20H, brs), 0.89 (3H, t,  $J = 6.9$  Hz, H-17').

#### **Supplementation of soybean oil into the fermentation broth of *E. coli* BL21(DE3)/pHW4**

*E. coli* BL21(DE3)/pHW4 was grown in 50 mL of LB with 50  $\mu$ g/mL kanamycin and induced by IPTG as described above except that 20 g/L soybean oil was supplied into the broth. The culture was extracted and analyzed by LC-MS as described above.

To isolate **7**, *E. coli* BL21(DE3)/pHW4 was grown in 200 mL of LB with 50  $\mu$ g/mL kanamycin and 20 g/L soybean oil. After fermentation, the broth was extracted and separated as described above. Compound **7** was found to be in fraction 5 after silica gel column chromatography. This fraction was further separated on HPLC with 74% acetonitrile-water. The peak at 30.7 min was collected to yield 2.1 mg of **7** in pure form.

Compound **7**: ESI-MS (+):  $[M+H]^+$   $m/z$  345.3;  $^1H$  NMR (500 MHz,  $CDCl_3$ ):  $\delta$  6.25 (2H, d,  $J = 1.8$  Hz, H-4 and H-6), 6.20 (1H, t,  $J = 1.8$  Hz, H-2), 5.34-5.41 (4H, m, H-8', H-9', H-11' and H-12'), 2.79 (2H, t,  $J = 15$  Hz, H-11'), 2.50 (2H, t,  $J = 7.5$  Hz, H-1'), 2.06 (4H, m, H-8' and H-14'), 1.57 (2H, brs, H-2'), 1.38 (approx. 14H, brs), 0.90 (3H, t,  $J = 6.9$  Hz, H-17').  $^{13}C$  NMR (125 MHz,  $CDCl_3$ ):  $\delta$  156.8 (C-1 and C-3), 146.1 (C-5),

130.2/130.1 (C-12'/C-8'), 128.0/127.9 (C-9'/C-11'), 107.9 (C-4 and C-6), 100.2 (C-2),  
35.9 (C-1'), 31.5 (C-15'), 29.7-28.7 (multiple carbon signals), 27.25/27.22 (C-7'/C-13'),  
25.6 (C-10'), 24.86 (C-2'), 22.6 (C-16'), 14.1 (C-17').

## **Heterologous expression of SsARS in *E. coli* and *S. cerevisiae* and product analysis**

*S. cerevisiae* BJ5464/pHW8 was grown in 50 mL of Synthetic Complete (SC)-Ura medium at 28°C to an OD<sub>600</sub> 1.0, and then 50 mL of YP medium (1% yeast extract and 2% peptone) was added. The broth was maintained at 28°C and 250 rpm for an additional 3 days. The fermentation broth was extracted and analyzed using the same method as described above.

## **Results**

### **Phylogenetic analysis of SsARS from *Shiraia* sp. Slf14**

Annotation of the genome sequencing data of *Shiraia* sp. Slf14 led to the discovery of a putative type III PKS SsARS in scaffold 13. BLAST analysis showed that SsARS is homologous to a number of type III PKSs. A phylogenetic analysis of SsARS and nine known type III PKSs from fungi, plant and bacteria revealed that SsARS falls within a clade of type III PKSs, and it forms a separate subclade together with *Neurospora crassa* ORAS (Funa et al. 2007) and *Sporotrichum laxum* Sl-PKS2 (Sun et al. 2016). As shown in Fig. 1, the horizontal dimension gives the extent of genetic change, and it was found that SsARS are more similar to Sl-PKS2 and ORAS. Both Sl-PKS2 and ORAS were previously reported to synthesize fatty acid-primed resorcinols, suggesting that SsARS is likely an alk(en)ylresorcinol synthase (ARS). Sequence alignment revealed that SsARS has 34% identity with *N. crassa* ORAS and 36% identity with *S.*

*laxum* SI-PKS2. SsARS maintains an identical CoA binding site with the reported type III PKSs and a conserved catalytic triad of Cys-153, His-308, and Asn-341 (numbering in SsARS).

### **Heterologous expression of SsARS in *E. coli* BL21(DE3) and characterization of the polyketide products**

The 1,278-bp intron-free cDNA of *Ssars* was synthesized from mRNA through reverse transcription. The gene was ligated to the expression vector pET28a to yield pHW4. *E. coli* BL21(DE3) was transformed with pHW4 and cultured in LB broth and induced with 200  $\mu$ M IPTG. After incubated at 28°C and 250 rpm condition overnight, the cells were collected and lysed for SDS-PAGE analysis. Compared to *E. coli* BL21(DE3) harboring the empty vector pET28a, a protein band at ~47 kDa was observed (Fig. 2A), suggesting that SsARS was successfully expressed in this heterologous host. We next examined whether this enzyme can generate any polyketide products in *E. coli*. The extracts of *E. coli* BL21(DE3)/pHW4 and *E. coli* BL21(DE3)/pET28a were analyzed by HPLC. As shown in Fig. 2B, BL21(DE3)/pHW4 produced a series of nonpolar products with retention times between 20 and 45 min (trace i), while the control with the empty vector did not have the same peaks. These six major products were numbered as **1-6**. Therefore, through SDS-PAGE and HPLC analysis, we confirmed that SsARS were functionally expressed in *E. coli* BL21(DE3).

### **Characterization of the six polyketide products from *E. coli* BL21(DE3)/pHW4**

To structurally characterize the six products, these compounds were isolated from the culture of *E. coli* BL21(DE3)/pHW4. Compounds **1-6** were obtained as yellowish oil.

These compounds were identified as alk(en)ylresorcinols with either an unsaturated or saturated aliphatic side chain (Fig. 2B) based on their UV, MS and NMR spectra and a comparison with previously published data (Suzuki et al. 1997 ). The molecular weights of **1-6** were determined to be 264, 290, 292, 318, 320 and 346 based on their  $[M+H]^+$  ion peaks at  $m/z$  265.2, 291.2, 293.2, 319.2, 321.4 and 347.3 in the ESI-MS (Fig. 3A), respectively. Compounds **1**, **3** and **5** were deduced to be alkylresorcinols with a saturated side chain, while **2**, **4** and **6** have an unsaturated side chain. The  $^1H$  NMR spectrum of **1** revealed two aromatic proton signals at  $\delta_H$  6.26 (2H, d,  $J = 1.8$  Hz, H-4 and H-6) and  $\delta_H$  6.22 (1H, t,  $J = 1.8$  Hz, H-2), suggesting that a resorcinol moiety is present in the structure. In addition, proton signals typical for a long aliphatic chain including the benzylic methylene at  $\delta_H$  2.49 (2H, t,  $J = 7.5$  Hz) and terminal methyl group at  $\delta_H$  0.90 (3H, t,  $J = 6.9$  Hz) were also found. Thus, **1** can be identified as 1,3-dihydroxyphenyl-5-undecane (Fig. 3B). The NMR data of **3** and **5** are highly similar to those of **1**, except that they have two and four more aliphatic proton signals, respectively. Accordingly, **3** was identified as 1,3-dihydroxyphenyl-5-tridecane and **5** was characterized as 1,3-dihydroxyphenyl-5-pentadecane (Fig. 3B).

There are two aromatic proton signals at  $\delta_H$  6.26 (2H, d,  $J = 1.8$  Hz, H-4 and H-6) and  $\delta_H$  6.20 (1H, t,  $J = 1.8$  Hz, H-2) in the  $^1H$  NMR spectrum of **2**, indicating that a resorcinol moiety is also present in the structure. Similar to **1**, proton signals for a long aliphatic chain including the benzylic methylene at  $\delta_H$  2.51 (2H, t,  $J = 7.5$  Hz) and terminal methyl group at  $\delta_H$  0.92 (3H, t,  $J = 6.9$  Hz) were also present. However, two overlapping olefinic protons at  $\delta_H$  5.38-5.36 indicated that there is a double bond in the

side chain. Thus, **2** was characterized as 1,3-dihydroxyphenyl-5-*cis*-6'-tridecene (Fig. 3B). Similarly, **4** and **6** were identified as resorcinol derivatives containing an unsaturated C<sub>15</sub> and C<sub>17</sub> side chain, named 1,3-dihydroxyphenyl-5-*cis*-8'-pentadecene and 1,3-dihydroxyphenyl-5-*cis*-10'-heptadecene, respectively (Fig. 3B). Compounds **2**, **4** and **6** were previously isolated and characterized by our group from an engineered *E. coli* strain that expressed Sl-pks2 from the biosynthetic pathway in spirolaxine in *Sporotrichum laxum* ATCC 15155 (Sun et al. 2016). Overall, the MS and <sup>1</sup>H NMR data of these known 5-alkenylresorcinols were consistent with our and others' reported data (Knodler et al. 2008; Sun et al. 2016; Suzuki et al. 1997 ).

#### **Effect of supplementation of exogenous fatty acids on the biosynthesis of 5-alk(en)ylresorcinols**

While the fatty acyl side chains of **1-6** were believed to be from the corresponding fatty acyl starter units caught by SsARS from the biosynthetic or degradative pathways within *E. coli* cells, we next tested how supplementation of exogenous fatty acids affects the product profile in *E. coli* BL21(DE3)/pHW4. Soybean oil is known to contain various fatty acids. To this end, 20 g/L soybean oil was supplied into the culture medium. The extract was analyzed by HPLC, and a new product **7** at 33.2 min was observed (Fig. 4A). This compound has a typical UV spectrum of alkenylresorcinols (Fig. 4B), suggesting that **7** is a derivative of alk(en)ylresorcinol.

Compound **7** was then isolated from the culture and subjected to MS and NMR analyses. The molecular weight of **7** was determined to be 344 based on the [M+H]<sup>+</sup> ion peak at *m/z* 345.3 (Fig. 4C), which is two mass units less than **6**, suggesting that an

additional double bond is present in the fatty acyl side chain of **7**. The  $^1\text{H}$  NMR spectrum showed the three aromatic protons at  $\delta_{\text{H}}$  6.25 (H-4 and H-6) and 6.20 (H-2), benzylic methylene protons at  $\delta_{\text{H}}$  2.50 (H-1'), and terminal  $\text{CH}_3$  signal at  $\delta_{\text{H}}$  0.90 (H-17'). Four overlapping olefinic proton signals were found at  $\delta_{\text{H}}$  5.34-5.41 and four corresponding carbon signals were found in the  $^{13}\text{C}$  NMR spectrum, including  $\delta_{\text{C}}$  130.2, 130.1, 128.0 and 127.9, confirming that there are two double bonds in the side chain. By comparing the  $^1\text{H}$  and  $^{13}\text{C}$  NMR data with those of linoleic acid (C18:2 *cis*, *cis*-9,12-octadecadienoic acid) (Bus and Frost 1976; Frost and Gunstone 1975; Glass and Dutton 1964), compound **7** was identified as 5-(8'*Z*,11'*Z*-heptadecadienyl)resorcinol.

We next compared the titers of compounds **1-7** in the cultures of *E. coli* BL21(DE3)/pHW4 with and without supplementation of soybean oil. As shown in Fig. 4E, it is apparent that supply of soybean oil increased the production of alk(en)ylresorcinols. Additionally, compound **7** was only produced when soybean oil was added into the fermentation medium, indicating that the exogenously supplied fatty acids can be used by SsARS to generate the polyketide products.

#### **Heterologous expression of SsARS in *S. cerevisiae* BJ5464 and characterization of the polyketide products**

The intron-free gene of *Ssars* was ligated to the YEpADH2p-URA3 vector to yield a yeast expression plasmid pHW8, which was then expressed in *S. cerevisiae* BJ5464. HPLC analysis of the fermentation broth of *S. cerevisiae* BJ5464 revealed that the same set of alk(en)ylresorcinols **1-6** were produced (trace i, Fig. 5), while the strain harboring the empty YEpADH2p-URA3 vector did not produce these products. This result

confirmed that SsARS can also be functionally reconstituted in yeast and synthesize the same products as *E. coli*.

## Discussion

*Shiraia* sp. Slf14 is an endophytic fungus from *H. serrata*. It was found to produce the medicinally important molecules (Yang et al. 2014). In this work, we found a putative type III PKS gene *Ssars* from the genome of this fungal strain and functionally expressed it in *E. coli* and *S. cerevisiae*. Expression of this enzyme in the two heterologous hosts led to the production of six major alk(en)ylresorcinols **1-6**, which unambiguously confirmed that SsARS is a type III PKS that specifically synthesizes resorcinol derivatives with flexible substrate specificity toward the fatty acyl starter units (Fig. 3B). The thioesters of common fatty acids in *E. coli* (Mejía et al. 1999; Shaw and Ingraham 1965 ) and yeast (Khalilova et al. 2016; Lamacka et al. 1998) are proposed to be used by SsARS as the starter units to synthesize **1-6**, with malonyl-CoA as the general substrate for polyketide chain elongation. Specifically, the thioesters of the unsaturated fatty acids including tetradecenoic acid (*cis*-7), hexadecenoic acid (*cis*-9), and octadecenoic acid (*cis*-11) were used by SsARS to generate **2**, **4** and **6**, while the thioesters of myristic acid (C14:0), palmitic acid (C16:0) and stearic acid (C18:0) acted as the starter units to afford **1**, **3** and **5**, respectively (Fig. 3B). Fatty acid composition of soybean oil of 18 varieties in the United States was reported to contain about 5% to 11% in linolenic acid (C18:3), 43% to 56% in linoleic acid (C18:2), 15% to 33% in oleic acid (C18:1), as well as 11% to 26% in saturated acids such as 4% stearic and 10% palmitic acids (Collins and Sedgwick 1959). These fatty acids

apparently have contributed to the production of the corresponding alk(en)ylresorcinols, leading to the increased production of **1-6** and the biosynthesis of an additional product **7** (from the major unsaturated fatty acid linoleic acid in soybean oil). This indicated that exogenous fatty acids can be used by the engineered *E. coli* strain to synthesize these alk(en)ylresorcinols.

Compared to many other natural products, alk(en)ylresorcinols are relatively rare in nature, with the main known sources being wheat, rye, barley, triticale (cereal grasses), the fruit of *Ginkgo biloba*, a mollusc and some species of bacteria. The concentrations of alk(en)ylresorcinols are normally between 300 and 1500 µg/g in wheat and rye (Ross et al. 2004). Due to their almost exclusive presence in the outer part of wheat and rye grains among commonly consumed foods (Ross et al. 2003), alk(en)ylresorcinols have been suggested and evaluated as biomarkers of whole grain wheat and rye intake in different populations (Biskup et al. 2016; Kyro et al. 2014; Magnusdottir et al. 2014; McKeown et al. 2010; Wierzbicka et al. 2015). A recent research revealed the dual nature of alk(en)ylresorcinols, direct antioxidant properties and the ability to indirectly regulate the activity of cellular antioxidative defense mechanisms (Gryazeva et al. 2015). Discovery and heterologous reconstitution of SsARS in this work provides an efficient way to produce desired alk(en)ylresorcinols by supplying corresponding fatty acids.

## **Conclusion**

In summary, we identified a new alk(en)ylresorcinol synthase SsARS from an endophytic fungal strain *Shiraia* sp. Slf14. This enzyme was functionally reconstituted



in both *E. coli* and *S. cerevisiae*. This fungal type III PKS showed relaxed substrate specificity toward endogenous and exogenous fatty acids. Depending on the availability of fatty acids in the cells or culture medium, different alk(en)ylresorcinols can be synthesized using this enzyme.

#### **Conflict of Interest**

The authors declare that they have no conflict of interest.

#### **Acknowledgments**

This work was supported by a Grant-In-Aid from the American Heart Association (16GRNT26430067), a National Science Foundation Award CHE-1429195, a Hunan Science and Technology Key Project (2014FJ1007), a grant from the National Natural Science Foundation of China (31460021), and the Hundred Talents Program of Hunan Province, China. R. Y. was supported by the Young Teacher Development Plan Visiting Scholar Special Funds of Jiangxi Province of China.

#### **References**

- Ali, S.M. and Olivo, M. 2002. Efficacy of hypocrellin pharmacokinetics in phototherapy. *Int. J. Oncol.* **21**, 1229–1237.
- Ali, S.M., Olivo, M., Yuen, G.Y., and Chee, S.K. 2001. Photodynamic-induced apoptosis of human nasopharyngeal carcinoma cells using hypocrellins. *Int. J. Oncol.* **19**, 633–64.
- Biskup, I., Kyrø, C., Marklund, M., Olsen, A., van Dam, R.M., Tjoenneland, A., Overvad, K., Lindahl, B., Johansson, I., and Landberg, R. 2016. Plasma alkylresorcinols, biomarkers of whole-grain wheat and rye intake, and risk of type

373 2 diabetes in Scandinavian men and women. *Am. J. Clin. Nutr.* **104**, 88–96.

374 **Bus, J. and Frost, D.J.** 1976. Determination of the positions of double bonds in  
 375 unsaturated fatty acids by carbon-13 and proton NMR spectrometry, pp. 343-350,  
 376 in, vol. 2. Raven.

377 **Collins, F.I. and Sedgwick, V.E.** 1959. Fatty acid composition of several varieties of  
 378 soybeans. *J. Am. Oil Chem. Soc.* **36**, 641–644.

379 **Damar, U., Gersner, R., Johnstone, J.T., Schachter, S., and Rotenberg, A.** 2017.  
 380 Huperzine A: A promising anticonvulsant, disease modifying, and memory  
 381 enhancing treatment option in Alzheimer's disease. *Med. Hypotheses* **99**, 57–62.

382 **Estey, E.P., Brown, K., Diwu, Z., Liu, J., Lown, J.W., Miller, G.G., Moore, R.B.,**  
 383 **Tulip, J., and McPhee, M.S.** 1996. Hypocrellins as photosensitizers for  
 384 photodynamic therapy: A screening evaluation and pharmacokinetic study.  
 385 *Cancer Chemother. Pharmacol.* **37**, 343–350.

386 **Frost, D.J. and Gunstone, F.D.** 1975. The PMR analysis of non-conjugated alkenoic  
 387 and alkynoic acids and esters. *Chem. Phys. Lipids* **15**, 53–85.

388 **Funa, N., Awakawa, T., and Horinouchi, S.** 2007. Pentaketide resorcylic acid  
 389 synthesis by Type III polyketide synthase from *Neurospora crassa*. *J. Biol. Chem.*  
 390 **282**, 14476–14481.

391 **Glass, C.A. and Dutton, H.J.** 1964. Determination of  $\beta$ -olefinic methyl groups in  
 392 esters of fatty acids by nuclear magnetic resonance. *Anal. Chem.* **36**, 2401–2404.

393 **Gryazeva, I.V., Davydova, O.K., and Deryabin, D.G.** 2015. Evaluation of the  
 394 potential of alkylresorcinols as superoxide anion scavengers and sox-regulon

modulators using nitroblue tetrazolium and bioluminescent cell-based assays.

*Cell. Mol. Biol. Lett.* **20**, 24–37.

**Kanda, N., Ishizaki, N., Inoue, N., Oshima, M., Handa, A., and Kitahara, T.** 1975.

DB-2073, a new alkylresorcinol antibiotic. I. Taxonomy, isolation, and

characterization. *J. Antibiot.* **28**, 935–942.

**Keatinge-Clay, A.T.** 2012. The structures of type I polyketide synthases. *Nat. Prod.*

*Rep.* **29**, 1050–1073.

**Khalilova, E.A., Kotenko, S.T., Islammagomedova, E.A., and Aliverdieva, D.A.**

2016. Comparative analysis of fatty acid composition in some *Saccharomyces*

*cerevisiae* strains. *Br. Microbiol. Res. J.*, **15**, 1–6.

**Kitahara, T. and Kanda, N.** 1975. DB-2073, a new alkylresorcinol antibiotic. II.

Chemical structure of DB-2073. *J. Antibiot.* **28**, 943–946.

**Knodler, M., Conrad, J., Wenzig, E.M., Bauer, R., Lacorn, M., Beifuss, U., Carle,**

**R., and Schieber, A.** 2008. Anti-inflammatory 5-(11'Z-heptadecenyl)- and 5-

(8'Z,11'Z-heptadecadienyl)-resorcinols from mango (*Mangifera indica* L.) peels.

*Phytochemistry* **69**, 988–993.

**Kyro, C., Olsen, A., Landberg, R., Skeie, G., Loft, S., Aman, P., Leenders, M., Dik,**

**V.K., Siersema, P.D., Pischon, T., et al.** 2014. Plasma alkylresorcinols,

biomarkers of whole-grain wheat and rye intake, and incidence of colorectal

cancer. *J. Natl. Cancer Inst.* **106**, djt352.

**Lamacka, M., Sajbidor, J., and Bohov, P.** 1998. Lipid isolation and fatty acid analysis

in *Saccharomyces cerevisiae*. Comparison of different methods. *Biotechnol. Tech.*

417 12, 621–625.

418 **Landberg, R., Kamal-Eldin, A., Andersson, A., Vessby, B., and Aaman, P.** 2008.

419 Alkylresorcinols as biomarkers of whole-grain wheat and rye intake: plasma

420 concentration and intake estimated from dietary records. *Am. J. Clin. Nutr.* **87**,

421 832–838.

422 **Landberg, R., Marklund, M., Kamal-Eldin, A., and Aman, P.** 2014. An update on

423 alkylresorcinols - Occurrence, bioavailability, bioactivity and utility as

424 biomarkers. *J. Funct. Foods* **7**, 77–89.

425 **Luis, A., Domingues, F., and Duarte, A.P.** 2016. Biological properties of plant-derived

426 alkylresorcinols: mini-review. *Mini. Rev. Med. Chem.* **16**, 851–854.

427 **Magnusdottir, O.K., Landberg, R., Gunnarsdottir, I., Cloetens, L., Akesson, B.,**

428 **Rosqvist, F., Schwab, U., Herzig, K.-H., Hukkanen, J., Savolainen, M.J., et al.**

429 2014. Whole grain rye intake, reflected by a biomarker, is associated with

430 favorable blood lipid outcomes in subjects with the metabolic syndrome - a

431 randomized study. *PLoS One* **9**, e110827/110821–e110827/110828.

432 **McKeown, N.M., Troy, L.M., Jacques, P.F., Hoffmann, U., O'Donnell, C.J., and**

433 **Fox, C.S.** 2010. Whole- and refined-grain intakes are differentially associated

434 with abdominal visceral and subcutaneous adiposity in healthy adults: the

435 Framingham heart study. *Am. J. Clin. Nutr.* **92**, 1165–1171.

436 **Mejía, R., Gómez-Eichelmann, M.C., and Fernández, M.S.** 1999. Fatty acid profile

437 of *Escherichia coli* during the heat-shock response. *IUBMB Life* **47**, 835–844.

438 **Ross, A.B., Kamal-Eldin, A., and Aman, P.** 2004. Dietary alkylresorcinols: absorption,

439 bioactivities, and possible use as biomarkers of whole-grain wheat- and rye-rich  
 440 foods. *Nutr. Rev.* **62**, 81–95.

441 **Ross, A.B., Shepherd, M.J., Schuepphaus, M., Sinclair, V., Alfaro, B., Kamal-Eldin,**  
 442 **A., and Aaman, P.** 2003. Alkylresorcinols in cereals and cereal products. *J. Agric.*  
 443 *Food Chem.* **51**, 4111–4118.

444 **Shaw, M.K., and Ingraham, J.L.** 1965. Fatty acid composition of *Escherichia coli*  
 445 as a possible controlling factor of the minimal growth temperature. *J. Bacteriol.*  
 446 **90**, 141–146.

447 **Sun, L., Wang, S., Zhang, S., Yu, D., Qin, Y., Huang, H., Wang, W., and Zhan, J.**  
 448 2016. Identification of a type III polyketide synthase involved in the biosynthesis  
 449 of spirolaxine. *Appl. Microbiol. Biotechnol.* **100**, 7103–7113.

450 **Suzuki, Y., Esumi, Y., Uramoto, M., Kono, Y., and Sakurai, A.** 1997. Structural  
 451 analyses of carbon chains in 5-alk(en)ylresorcinols of rye and wheat whole flour  
 452 by tandem mass spectrometry. *Biosci. Biotechnol. Biochem.* **61**, 480–486.

453 **Wierzbicka, R., Wu, H., Franek, M., Kamal-Eldin, A., and Landberg, R.** 2015.  
 454 Determination of alkylresorcinols and their metabolites in biological samples by  
 455 gas chromatography-mass spectrometry. *J. Chromatogr. B: Anal. Technol.*  
 456 *Biomed. Life Sci.* **1000**, 120–129.

457 **Yang, H., Wang, Y., Zhang, Z., Yan, R., and Zhu, D.** 2014. Whole-genome shotgun  
 458 assembly and analysis of the genome of *Shiraia* sp. Strain Slf14, a novel  
 459 endophytic fungus producing huperzine A and hypocrellin A. *Genome Announc.*  
 460 **2**, e00011-14.

461 **Yu, D., Xu, F., Zeng, J., and Zhan, J.** 2012. Type III polyketide synthases in natural  
462 product biosynthesis. *IUBMB Life* **64**, 285–295.

463 **Zhan, J.** 2009. Biosynthesis of bacterial aromatic polyketides. *Curr. Top. Med. Chem.*  
464 **9**, 1598–1610.

465 **Zhu, D., Wang, J., Zeng, Q., Zhang, Z., and Yan, R.** 2010. A novel endophytic  
466 Huperzine A-producing fungus, *Shiraia* sp. Slf14, isolated from *Huperzia serrata*.  
467 *J. Appl. Microbiol.* **109**, 1469–1478.

484 Figure legends

485 **Fig. 1. A phylogenetic tree of SsARS and nine known type III PKSs.** The nine type  
486 III PKSs include CsyA from *Aspergillus oryzae*, GenBank accession no. BAD97390.1;  
487 CsyB from *A. oryzae*, GenBank accession no. BAD97391.1; CsyC from *A. oryzae*,  
488 GenBank accession no. BAD97392.1; CsyD from *A. oryzae*, GenBank accession  
489 no. BAD97394.1; DpgA from *Streptomyces toyocaensis*, GenBank accession  
490 no. Q8KLK5.1; OKS from *Aloe arborescens*, GenBank accession no. Q3L7F5.1;  
491 StTS from *Streptomyces toxytricini*, GenBank accession no. AEO44526.1;  
492 ORAS from *Neurospora crassa*, GenBank accession no. XP\_960427.1; and Sl-  
493 PKS2 from *Sporotrichum laxum*, GenBank accession no. AMW87980.1.

494 **Fig. 2. Heterologous reconstitution of SsARS in *E. coli* BL21(DE3).** (A) SDS-PAGE  
495 analysis of the expression of SsARS in *E. coli* BL21(DE3). (B) HPLC analysis (at 210  
496 nm) of the polyketide metabolites produced by *E. coli* BL21(DE3)/pHW4. (i) Extract  
497 of *E. coli* BL21(DE3)/pHW4; (ii) Extract of *E. coli* BL21(DE3)/pET28a.

498 **Fig. 3. Structural characterization of 1-6.** (A) The ESI-MS (+) spectra of 1-6. (B)  
499 Structures of compounds 1-6 and their biosynthetic pathway.

500 **Fig. 4. Precursor-directed biosynthesis of an alkylresorcinol product with soybean oil.**  
501 (A) HPLC analysis (at 210 nm) of the extract of *E. coli* BL21(DE3)/pHW4 in LB  
502 supplemented with 2% soybean oil. (B) UV spectrum of 7. (C) ESI-MS (+) spectrum  
503 of 7. (D) Structure of 7. (E) The titers of 1-7 in the cultures of *E. coli* BL21(DE3)/pHW4  
504 in LB with or without soybean oil. Data were from three experiments and are shown as  
505 mean $\pm$ SD.

**Fig. 5.** HPLC analysis (at 210 nm) of the products synthesized by SsARS in the yeast.

(i) *S. cerevisiae* BJ5464/pHW8; (ii) *S. cerevisiae* BJ5464/YEpADH2p-URA3 (vector control).



Fig. 1.

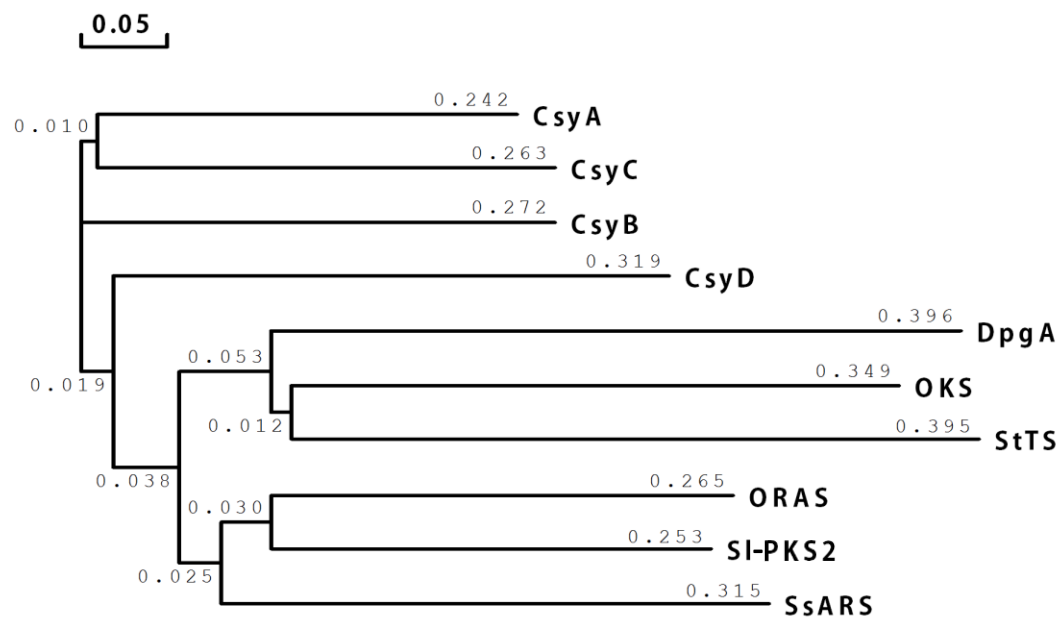


Fig. 2.

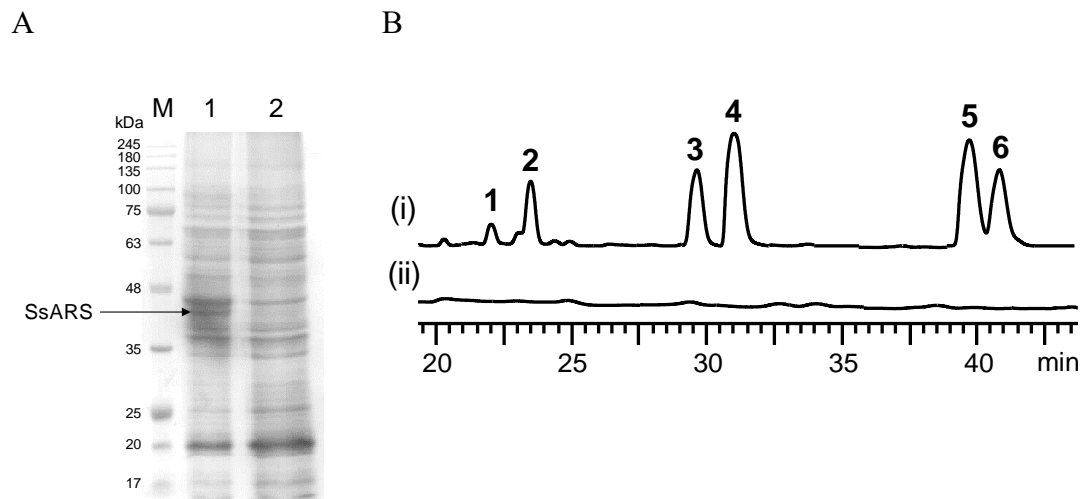


Fig. 3.

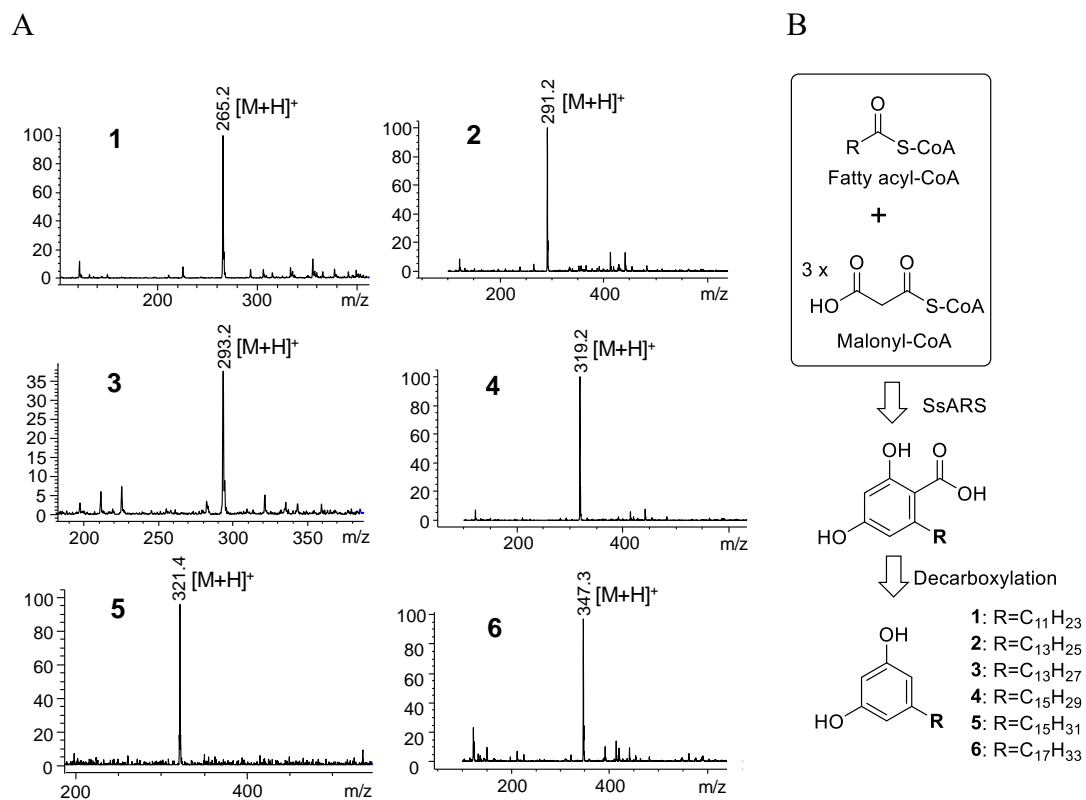
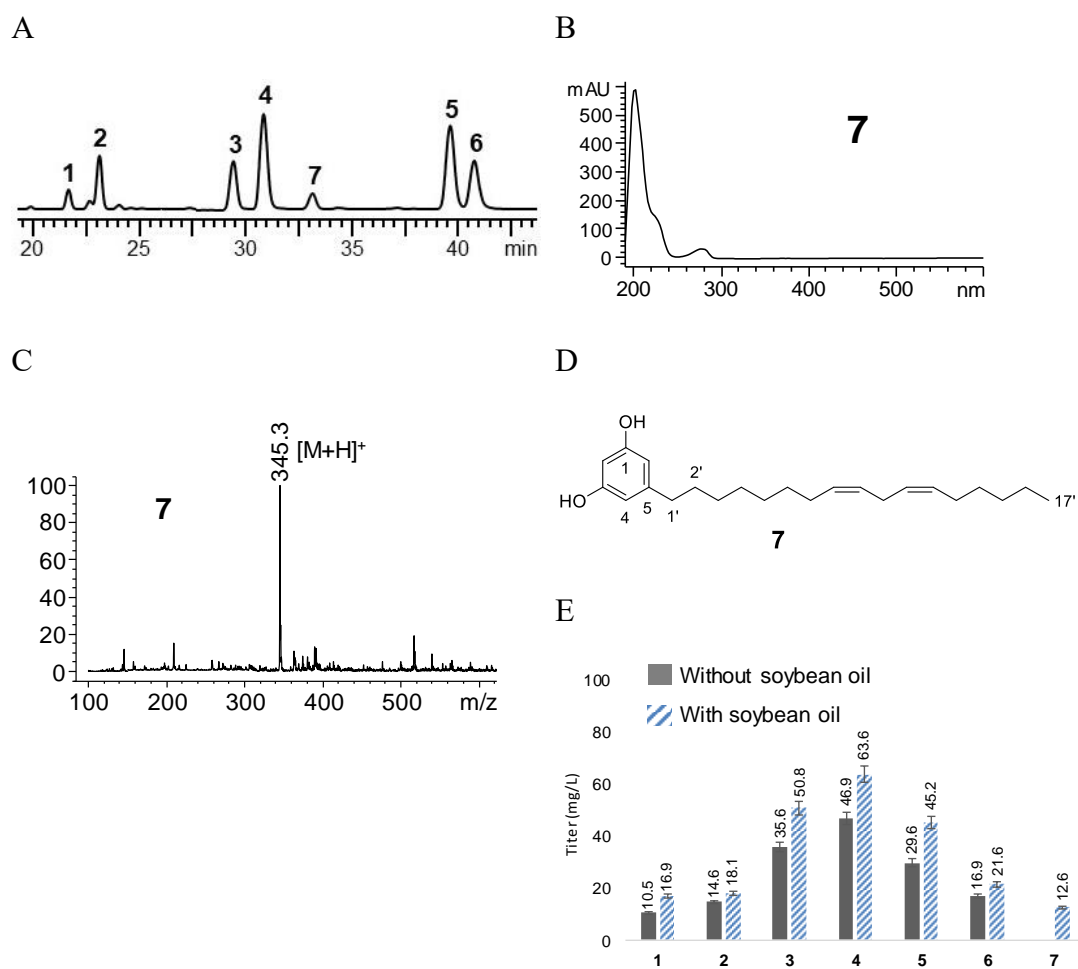


Fig. 4.

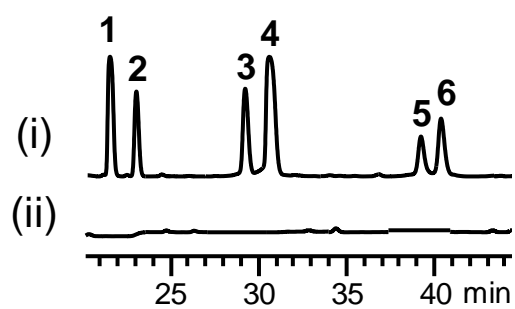


582

583 Fig. 5.

584

585



586

587